R E P O R T R E S U M E S

ED 003 911 48

A TYPOLOGICAL CLASSIFICATION OF DICTIONARIES ON THE BASIS OF DISTINCTIVE FEATURES.

BY- MALKIEL, YAKOV
INDIANA UNIV., BLOOMINGTON

REPORT NUMBER NDEA-VI-16-4 PUB DATE 60
EDRS PRICE MF-$0.09 HC-$1.20 30P.

DESCRIPTORS- #DICTIONARIES, #BIBLIOGRAPHY, #HISTORICAL REVIEWS, #CONFERENCES, #LEXICOGRAPHY, BLOOMINGTON, INDIANA

A SET OF BROAD CLASSIFICATORY CRITERIA WAS PRESENTED FOR USE IN LEXICOGRAPHICAL RESEARCH. THESE CRITERIA CONSISTED OF FUNDAMENTALS FOR CLASSIFYING MATERIAL BY RANGE, PERSPECTIVE, AND PRESENTATION. THIS REPORT IS ONE OF A SERIES OF 13 PAPERS PRESENTED AT A CONFERENCE ON LEXICOGRAPHY, INDIANA UNIVERSITY, NOVEMBER 11-12, 1960. (JH)

ED003911

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

16-4
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

A Typological Classification of Dictionaries on the Basis of Distinctive Features

Yakov Malkiel, University of California, Berkeley

1. Background

The present paper embodies a report on an experiment undertaken some two years ago. The results of this specific experiment may or may not lend themselves to generalization. From the mid-'forties, I had engaged for fifteen years in various studies of Romance etymology and derivation, with a clearly discernible emphasis on Hispanic material. These researches were based in part on the perusal of original texts, but, inescapably, a wide range of dictionaries had been tapped for miscellaneous bits of information. There exists a tradition among lexicologists to offer to their fellow-workers, besides the genuine fruits of their researchos, also an inventory of the tools employed. Thus, W. von Wartburg's monumental Französisches Etymologisches Wörterbuch, slanted in the direction of dialectology, was accompanied by a bibliography of patois vocabularies. Instead of preparing, along traditional lines, a similar Hispanic bibliography, containing items either arranged in a straight alphabetical sequence or subdivided into several regional groups--say, Asturian, Andalusian, Mexican--, I began asking myself whether a pervasive classification by lexicographic genres was at all feasible. A rough break-down of this kind was, of course, nothing new, since numerous research-library catalogues make it a point to distinguish between mono-, bi-, tri-linguial and polyglot dictionaries and likewise set apart encyclopedic, historical, comparative, and etymological reference words and surely many other sharply delineated varieties, especially those appealing to a specific kind of reader. As is frequently the case with classification, the core of the difficulty seemed to lie in the overlapping of the categories hastily established on the basis of first impressions. To adduce but one example:

aside from straight etymological dictionaries, geared to the one purpose of supplying information on word origins (on various levels of scholarly sophistication), there exist all-purpose monolingual dictionaries and, far less commonly, bilingual dictionaries which, in response to the layman's curiosity, also carry stray bits of etymological information, representing, in a way, mixed genres of lexicography. In simplified listing, should such transitional types be altogether disregarded and the pure types alone be placed in the focus of attention? Such an arbitrary decision would tend to curtail unduly the available material. The alternative would be to do justice to as many titles as possible, on condition that either the number of categories be radically increased or that each item, i.e., each lexicographic venture, instead of being treated as an inseparable unit, be decomposed into a number of salient features allowing of a large number of combinations. The obvious model that I had in mind was the separation of a speech sound into a number of distinctive features, an analysis originally practiced with particular skill by the school of European phonologists, but at present widely adopted on both sides of the Atlantic. Was it legitimate to regard a reference book as a bundle of characteristic features which could be reassembled in some kind of arrangement that could be expressed in a simple formula? The article in question stopped short of this last step of providing actual formulas, but was designed to lead the reader to the precise point from which such a succinct formulaic labeling could be freely envisaged.[1]

For this particular experiment, limited to a single language and involving, all told, not much more that five hundred items, the distinctive features were chosen so as to satisfy two fundamental requirements: range and neatness of contour. First, the greatest possible number of the vocabularies and glossaries collected at random were to lend themselves

to this type of composite characterization. Second, any degree of redundancy in the delimitation of distinctive features was to be scrupulously avoided, each being selected on the basis of its discreteness, i.e., independence of the others. The existence of tendential convergences and affinities was, from the outset, taken for granted, but every feature, to qualify for inclusion, was to preserve its unmistakable identity under any circumstances.

2. A set of broad classificatory criteria

These aims, it seems, call for the establishment of the following three classificatory criteria, each subject to further subdivision:

1. Classification by range;
2. Classification by perspective;
3. Classification by presentation.

Range, i.e. the volume and spread of the material assembled, represents, from the layman's and the specialist's viewpoint alike, the most obvious criterion; it is also the most objective, involving by definition tangibles alone. Further subdivision is possible, according as to whether attention is focused on (a) density of entries, (b) number of languages covered or resorted to, (c) degree of concentration on strictly lexical data (at the expense of realia, proper names, and the like).

Perspective: under this label it is advantageous to subsume several broad directions in which bundles of lexical facts are apt to be projected, irrespective of the scope of each vocabulary and of the favored style of presentation. Perspective essentially involves the deliberate or semi-naive attitude of the collector toward the chosen alice of material, i.e., the type of curiosity that drives him in the first place to delimit and launch his project. The specific attitude is best determined if one selects the proper components of the ensemble from a number of oppositions,

binary or triadic: (a) with regard to the time axis, the outlook may be historical (dynamic) or synchronic (static); (b) with regard to the basic arrangement, the sequence may be conventional (alphabetic), semantic (ordering by "parts of speech" or provinces of life), or entirely arbitrary (chaotic); (c) with regard to avowed purpose, the prevalent tone may be objective, preceptive (also, viewed from the opposite pole, prohibitive), or jocular. In theory, any vocabulary may be scrutinized from all three angles, and the combination of the three separate analyses determines the composite perspective. On a higher plane of abstraction one may attempt to establish certain steady or tendential relationships between subclasses within the three categories of temporal projection, sequence of items, and purpose of enterprise.

Presentation: if operating with "perspectives" helped us account for the broad policies, "presentation" serves as a convenient term for subsuming narrower preferences, such as typographic style, use of special symbols and abbreviations, inclusion of incidental grammatical remarks, latitutde of definition, volume of verbal documentation and of graphic illustration (if any), and many similar externals, In other words, perspective concerns the lexicographer's logistics and strategy, while presentation involves his tactics. To achieve simplification, analysis may be reduced to the discussion of four salient points: (a) definition; (b) exemplification; (c) graphic illustrations (including maps); (d) special features (localization in territorial terms, on the social scale, or along the axis of "affectivity"; marking of pronunciation).

Ideally, a formal bibliography or a research-library catalogue may well be based on this (or any improved) set of classificatory principles. While conventional bibliographies and catalogues adequately describe the externals of a publication (formal, number of pages, place and date of

appearance, etc.), a typological repertory may effectively convey an analytical bird's-eye view of the inner configuration of each item (short of value judgments): its spatial and temporal range, its degree of selectivity, the interplay of its three basic perspectives (distance from the historical axis, schema of arrangement, and level of discourse), plus its editorial and typographic techniques of presentation.

3. Some special conditions

Last year's original monograph underlying the present paper supplied only a few particularly representative instances of a given variety or subvariety, sometimes out of more than twenty available illustrations. As long as the classificatory schema is valid, expanding such a skeletal documentation to one of fuller size or of relative completeness remains a matter of space budgeting. Little more than a vague hint of the preliminary bibliographic spadework can be provided in the present epitome.

What marks an item for inclusion in a typological survey is less the intrinsic quality of craftsmanship than the novelty of the architectural design. In extreme cases, an elaborate work, structured along conventional lines, may fall short of acceptance, if it enriches the existing models by no truly original device or daring combination of familiar devices. Conversely, a departure from earlier practice--even if deemed by hindsight a failure or, at best, a controversial accomplishment--deserves consideration, as an attempt at innovation. At its most austere, a typological classification should steer clear of value judgments; in practice, succinct appraisals amounting to merely parenthetic remarks may be condoned, as long as they do not interfere with actual analytical operations.

Aside from its intrinsic significance and charm, the typological approach may produce a wholesome controlling effect: It is apt to disclose noteworthy gaps, identifying combinations of features not yet tried out,

and may thus open up avenues of approach not only to all sorts of impracticable oddities, but also to rewarding possibilities so far overlooked. In chosen instances one can point out how similar lacunae have been successfully filled by fellow-workers concerned with other languages.

Typological sketches are not meant to replace the time-honored genre of formal bibliographies, which are, by and large, more circumstantial, silhouette with equal sharpness each work captured by the observer's lens, and lend themselves smoothly to evaluative comments. Typology, being the more abstract mode of classification, simply acts as a corrective of straight bibliographic records, breaking loose of anecdotal detail and of the all too often obtrusive "human element". Jointly, bibliography and typology reflect the dual focus of linguistics on individual facts of language, viewed in their undistorted richness and uniqueness, and on the underlying broader patterns.

4. Classification by range

This approach, we recall, allows of further subdivision, according to the variable emphasis on (a) density of entries, (b) number of languages, and (c) extent of concentration on purely lexical data.

(a) Measurement of density. Though ideally a dictionary's relative density, being a quantitative feature, should be subject to precise measurement, there exist in practice almost insurmountable obstacles to a neat statistical tabulation. It is, of course, possible to compute the actual number of entries, and some publishers flaunt it in a subtitle or in commercial advertisement (others claim credit not for the total coverage, but for specific accretions, choosing some earlier authority as an appropriate frame of reference). The figures thus obtained, though accurate in isolation, lend themselves to no meaningful comparison, in the absence of generally accepted norms for the selection and grouping of entries.

Thus, a few reputable lexicographers, for the sake of maximum compactness, consolidate homophones, subsuming under a single entry two or more words treated separately by such (more conservative) authors as prefer to heed semantic or etymological considerations.

In the case of a living and dynamically changing language, the primary hindrances to a valid comparison of coverage are: (a) the essential instability of its lexicon, a condition precluding the establishment of any definitive repository based on exhaustive excerpts from a finite corpus of texts; (b) the, strictly speaking, incompatible scopes of even the best lexicographic projects so far launched, as a result of sharply divergent objectives and techniques. Among the compilers of dictionaries few have bothered to lay down binding norms or to abide by any. Numerous dictionaries allow the inclusion of archaisms, whether or not marked as such; only a few workers resolutely attempt to weed out obsolete terms by starting from entirely fresh collections of raw data or settle on a compromise by making carefully preëstablished concessions to the dominant preference (e.g., by respecting the usage of a few classics, explicitly identified). As regards coverage in breadth, it seems inadmissible, then, to maintain that a given dictionary contains, say, 80% of the total lexicon, as long as that total remains undefined, and it is hazardous to assert that one dictionary is more comprehensive than another, unless they are modeled on exactly the same pattern. To compound the difficulty, lexical range is concomitantly determined by coverage in depth, i.e., by the wealth of recorded meanings (including, rightly or wrongly, contextual connotations) and congealed sequences ("idioms"). Since an objective, absolute classification of meaning has so far remained an unattainable goal, much depends on the intended degree of elaborateness and, in a bilingual venture, also on the distance between the two languages at issue: for practical purposes, a

Spanish-Portuguese dictionary need not be as detailed on the phraseological side as, say, its Spanish-Aztec counterpart. Under these circumstances, it may often be wiser to dispense altogether with statistics and to rely on rough impressionistic appraisals for their provisional usefulness. One paradoxical comment on monumentality in this field: frequently, the most voluminous collections must receive a conspicuously low rating on grounds other than that of sheer size.

At the opposite extreme of the continuum, one encounters exceedingly brief and casual word-lists: collections of medieval glosses (some of them very meager and not even organized as miniature vocabularies), miscellaneous records of dialect forms, countless short arrays of words pertaining to ultramodern slang or to certain professional jargons of striking, sometimes perverse, appeal to readers at large. Most of these lists, scattered over obscure or ephemeral periodicals, are doomed to remain practically unavailable for ordinary research purposes. Other short word-lists difficult of access to the uninitiated include the vocabularies--some of them gems of neat workmanship--appended to collegiate editions of literary masterpieces in classical and modern foreign languages and, on the tradebook market, glossaries accompanying contemporary novels and short stories written in a national language, but, for the ordinary reader's comfort at least, excessively spiced with regional flavor. To these nuggets of knowledge one may add those extra-brief word-lists (sometimes reduced to a single page) that form part of the typical "academic" editions of classical and medieval texts. There exist trustworthy catalogues of such editions, both mono- and bi-lingual, but very few and, at that, fragmentary master-lists of the items actually contained in these minimum lists.

We have so far worked on the assumption that dictionaries represent

variously scaled abridgments of an ideally complete lexicographic record. This premise requires a corrective: not a few compilations include words that have never existed in the reality of living speech or of written usuage. This excess baggage embraces, first of all, isolated ghost words --old misprints or spurious readings inadvertently carried over from one dictionary into another. Such blemishes modern workers, at the cost of hard labor, endeavor to ferret out one by one. Far more numerous and, worse, more insidious are the latent words smuggled in deliberately. In languages possessing an arsenal of productive suffixes (such as Slavic or Romance) a lexicographer can quite unobtrusively manufacture, by the hundreds and thousands, derivatives of his own making. These formations are readily understandable and give the appearance of authenticity; the truth is that speakers, for some reason or other, have not bothered to activate on the same scale all these grammatical potentialities. It is difficult to detect such camouflaged illicit entries, especially in view of legitimate border-line cases (nonces); an unduly geometric design of representative word families, a dictionary's too heavy saturation with certain recurrent derivational schemes at once arouses the suspicion of an experienced reader.

The average dictionary not only tends to overlook differences in chronological levels, but, as a rule, represents a medley of diverse social and regional dialects. Full-scale lexical treatment of a strictly defined social dialect has so far been relatively infrequent; concomitantly, however, the social category of speakers may be implied in the researcher's choice of a regional sector (rural versus urban speech). In inquiries concerned with peculiar linguistic milieus (immigrants, "bohemians", and the like), the threads of regional and of social dialects are inextricably interwoven.

At this point, the original monograph examines in considerable detail the patterns of dialect vocabularies and those of philological glossaries.

The unadorned dialect vocabulary must be distinguished from monographs balancing lexicography with grammar and with specimens of texts, also from word indexes attached to or implied in grammatical analyses. The dialect area selected as a unit may vary from a hamlet to a continent, the controlling factors being the diversification of speech, the configuration of the terrain, the pattern of political allegiances, and the availability of field-workers and informants. As research grows more and more specialized, master dictionaries for dialect groups become an urgent desideratum. A complementary technique for gauging the scope of a regional vocabulary based on field-work is to determine the number of interviewers and interviewees. Most regional vocabularies, instead of affording a total view of the chosen lexicon, are satisfied with recording deviations from standard usage, like so many tidbits; a few subdivide them further into archaisms, broad and narrow localisms, foreignisms, etc. A special, neatly detachable strain in a given regional variety may be the object of a separate vocabulary. This is true in particular of those monographs which concern themselves with autochthonous Indian words absorbed by a national variety of American Spanish, e.g., the Araukan components of Chilean.

The range of a philological vocabulary may be delimited by the distance between points selected along the time axis. The historical dictionary of the future may dissolve into a series of chronologically bounded vocabularies reflecting a roughly fixed number of superimposed layers. One finds a variety of narrower scopes in exegetic vocabularies, assembled around the usages of individual authors, texts, or genres--some of them aiming at exhaustiveness (concordances), while others are overtly and explicitly, or tacity and loosely, selective. The compiler's readiness to do full or partial justice to the manuscript tradition and the inclusion of supporting evidence, in varying dosages, from miscellaneous collateral sources, yield two further yardsticks for the measurement of lexicographic scope. Though dialectology stresses space, while philological analysis entails movements along the time axis, the two approaches are not mutually exclusive, since artistic works may display adequate and not infrequently unique specimens of speech dialectally tinged--inviting glossaries straddling the two disciplines.

The density of a vocabulary may be determined by the degree of commonness (in statistical terms, frequency) of the words to be included --a distinctly modern criterion. In certain lines of practical work (preparation of graded elementary textbooks, dilution of fine literature for classroom use) the need has arisen for precisely circumscribed core vocabularies. Some workers distinguish between the coördinates of "range" and "frequency", for which one may substitute breadth and depth of penetration. Aside from statistically slanted word lists there exist at present comparable tabulations of phrases (idioms) and syntactical constructions; from all such storehouses of data one may extract direct or, at least,

indirect information on this particular aspect of the lexicon. Printed sources may be freely supplemented or even replaced by samples of spontaneous or controlled speech. One may subdivide the body of the vocabulary into several layers, on the basis of increasingly higher frequency; one worker (1941) discriminates between four such groups, recordiigg them both separately and in a master-list in which raised numerals act as classifiers and noteworthy transparent derivatives not caught in the meshes of the statistical network are added as a bonus.

The various formats that have become the stock-in-trade of the publishing business specializing in dictionaries frequently offer excellent outlets for the varying patterns of amplitude selected by the lexicographers --or by the masterminds in charge of their projects.

(b) Number of languages involved. A second way of classifying dictionaries by range is to rate them by the number of languages covered or resorted to. Does the use of a second language entail no more than a special type of definitions, namely translations, and does the contrast between mono- and bi-lingual works consequently involve a matter of presentation rather than of scope? The answer to this question is in the negative; not only is the argument hardly applicable to pluri-lingual dictionaries, which patently mark an increase in coverage, not a refinement in explanatory technique, but worse, it discounts the fact that numerous bilingual dictionaries are bipartite, hence geared to serving two separate needs measurable in terms of sheer volume.

The monolingual dictionary constitutes a fairly recent genre: even the earliest editions of many tone-setting Academy dictionaries, which one inclines to regard as monolingual almost by definition, made it a point to translate all entries into Latin. In a conservative country like Spain, the unadulterated monolingual dictionary is a creation of the

nineteenth century. By modern standards, the only admissible residue of Latin in any such reference work would be the identification of vernacular names of, say, plants and animals by the tags assigned to them in international scientific nomenclature.

The bilingual dictionary (or, if it happens to be short, glossary) is the vocabulary par excellence, associated from time immemorial with casual exposure to, or systematic training in, foreign languages, dead or living. In Western civilization, which is not the sole inventor of this genre, rudimentary glossaries (Greek-Latin, Latin-Germanic, Latin-Romance) are traceable to late Antiquity and the early Middle Ages; throughout, the preponderance of Latin, either as the tool language providing the "glosses", or as the language requiring itself a measure of exegesis, is quite evident. It took the tremendous shifts of the sixteenth century to creat a large-scale demand for bilingual reference works involving either two modern western languages (say, Spanish and Italian, or Spanish and English), or one western and one "exotic" language, the latter typically one of special concern to colonizers and missionaries (e.g. Granadine Arabic, Aztec, Maya). The only recent additions to this stock of patterns are dictionaries prepared in the "western" style, but serving to contrast two "exotic" languages --of which one, for instance, may very well be Japanese; and a sprinkling of dictionaries pairing off a major "natural" with some "artificial" or "auxiliary" language, e.g. Esperanto, Ido (or Reformed Esperanto), and Volapük.

Many, if not most, bilingual dictionaries are bipartite, irrespective of size and quality. Occasionally (as in the case of Slabý-Grossmann's German-Spanish venture) the two parts were produced by different compilers, each starting with a clean slate, and, as a result, show unintentional discrepancies in bulk and craftsmanship. Then again, in some instances the

two sections deliberately have not been planned on the same footing, the need for translations from the foreign into the native language being apparently deemed far more urgent than translation in the reverse direction. This consideration favors the coupling of a full-sized "dictionary" and a meager "vocabulary" typically one fourth or one fifth its size, which in reality is little more than a makeshift index. This asymmetric architectural design would be inappropriate where the two languages are equal partners, as regards social prestige and actual use (say, English and Russian); it is marginally acceptable if some kind of hierarchy has crystallized, the dead language normally surpassing in importance its living counterpart as an object of lexicographic curiosity, since few persons would care to translate modern texts <u>into</u> Latin or Old Norse, while dialect speech and cant, for similar reasons, carry greater weight as exhibits than the corresponding literary tongues, used preëminently as "tool languages".

At least two categories of bilingual dictionaries are seldom bipartite. Temporarily, this restriction applied to those compiled by proselytizing clergymen for overseas use. Then again, dialect vocabularies and glossaries of medieval words—if at all classifiable as bilingual—almost never contain so much as an index of meanings cast in the modern language chosen as the frame of reference, let alone a full counterview of lexical relationships, to the lasting detriment of advanced research.

A trilingual dictionary may be the outgrowth of a bilingual prototype, whether the author (or a reviser) himself arranges for the expansion or a plagiarist enlarges upon a chosen model. Frequently a preëxistent cultural or political climate favors such a necessarily high-aiming venture. Thus the dual foundation of classical learning accounts for such dictionaries as involve Greek, Latin, and a modern vernacular. Since the interest of Spaniards, Frenchmen, and the Dutch-speaking natives for a while converged

on Flanders, it is small wonder that, in response to this demand, a trilingual dictionary should have been engineered at the critical point (1640). Two important French-Spanish-Basque dictionaries, M. de Larramendi's (1745) and R. M. de Azkue's (1905-06), have their roots in the geographic position of the Euskarian territory, astride the Pyrenees. In another context, one may find a trilingual dictionary at the intersection of two or three cultural traditions; cf. the obvious historical pattern underlying F. Cañes' voluminous Diccionario español-latino-arábigo (1787). As a rule, trilingual dictionaries, chiefly for reasons of economy, are unidirectional, i.e., each offers maximum usefulness only to that segment of consulters which has mastered the privileged among the three languages (e.g., Spanish in the case of Cañes). Where differently architectured works of this type bear on the same subject, the expert is best served by combining imaginatively these sources of information.

Theoretically the quadrilingual dictionary leaves room for a vastly increased variety of sequences and combinations, but in practice few of these mathematical possibilities have been exploited on account of the comparative rarity of the genre. The kernel of Howell's genteel Lexicon tetraglotton (1660) is a unidirectional dictionary providing for the translation of each English entry into French, Italian, and Spanish. Terreros y Pando's life-work contains three volumes of consecutive Spanish entries translated into French, Spanish, and Italian, and, in a fourth volume, three alphabetic indexes in the reverse direction, each with separate paging.

In contrast to the polyglot Bible, a milestone in the growth of Humanism, the multilingual dictionary, as a scholarly institution, has failed to exert any lasting influence on European (still less on American) intellectual life. Compilations of this kind have usually been manufactured either as saving devices (to economize time, effort, production costs,

shelving space) or as veritable tours de force, for the sake of the bizarre effects. Of late, the best-known among them have tended to involve from five to ten languages; the arrangement offers almost limitless possibilities.

On the whole, dictionaries contrasting more than two languages have failed to maintain the position of eminence which the intellectual élite assigned to them in the 17th and 18th centuries, as works of scholarship and even of art. In the contemporary world, they have been relegated to the unattractive role of tools in strictly commercial and industrial relations, quite unexciting on the linguistic side. The need for these glossaries of technical terms may soon altogether recede as a result of the invention of electronic translating devices. In a way, the steady decline of the polyglot dictionary has been, paradoxically, a direct consequence of increased lexicographic sophistication: if the semantic, syntactic, and phraseological details which a fastidious reader nowadays demands of his mono-or bi-lingual dictionary were furnished on this grandiose scale, the resulting increase in bulk would at once defeat the chief purpose of lexical telescoping.

(c) Extent of concentration on lexical data. Much as biographical dictionary and straight bibliography for centuries tended to represent a single undifferentiated genre, so dictionary and encyclopedia were not always recognized as guides to discrete provinces of knowledge. The infiltration of any encyclopedia-style data is recognizable (a) by the inclusion of proper names, sometimes accompanied by biographic vignettes and profuse geographic descriptions; (b) by a prodigality of comments on ordinary words that seems far in excess of the need for sober definition.

Encyclopedic entries may be relegated to a supplement or a string of supplements (gazetteer, etc.), as has become the dominant fashion in recent decades. Also, they may be intercalated, so as to form part of the same unique alphabetical succession of items as the ordinary words. Finally, the

trained observer may be in a position to grasp them obliquely, through internal analysis of the definitions. In the first eventuality, there subsists a clear boundary between purely lexical and other bits of information; in the second, the frontier-line is less sharply perceptible; in the third, these extraneous, non-lexical ingredients are submerged beneath the unruffled surface of a deceptively balanced and unified dictionary.

Where the publisher furnishes supplements, by way of bonuses to prospective buyers, one finds a motley array filling needs at present preferably met by separate books (rosters of proper names, guides to pronunciation, grammar, and conversation, collections of proverbs, miniature rhyming dictionaries, lists of synonyms, corrections of misprints, and the like). It has likewise become increasingly customary to provide a separate list of toponyms and anthroponyms in glossaries to medieval texts.

The larding of dictionaries with non-lexical elements, presented as matters of sober fact or in a genial vein, became fashionable toward 1700. In some of its key points (Paris, Barcelona) the book trade, for two long centuries, contined stuffing dictionaries with shreds of information which no discriminating linguist would, but the less knowledgeable public at large frequently does, expect a lexicon to provide into the bargain, as it were--through interfiling, in the compiler's workshop, of authentic word cards and brief excerpts from encyclopedias. This hybrid genre has been mostly known under the name of "encyclopedic dictionary". The same publishing centers have tolerated multi-volume encyclopedias shot through, in defiance of common sense, with genuinely lexical entries.

The least easily detectable dilution of a bi- or tri-lingual dictionary is achieved by the inclusion of such parenthetic information as is linguistically immaterial. The ideally tight dictionary adds comments to translations only under three conditions: (a) where ambiguity might result from maximum

concision, as in the case of homonyms; (b) where circumlocution is inherently in order, on account of a physically, socially, or linguistically conditioned lack of nearly perfect lexical equivalents; (c) where foreseeable grammatical complications may arise. Any other remarks should be rated as redundant. At certain periods, this ideal of stringency was not yet--or no longer--appreciated. Thus, in the Renaissance dictionary, the two halves of each succinct correspondence were of approximately equal length, except where a lexicographer bothered to adduce "authorities" insupport of the posited equations. Two centuries later, at the height of the "encyclopedic vogue," the entries, as a rule, remained brief, but the glosses tended to become more and more inflated.

5. Classification by perspective

The three basic perspectives, as defined at the outset, involve (a) the fundamental dimension (diachronism versus synchronism), (b) the basic arrangement of entries (conventional, semantic, or arbitrary), (c) the level of tone. Accordingly, the typologist must arrange for three separate analyses.

(a) The fundamental dimension. The ideal synchronic dictionary would be one least contaminated by acknowledged or, worse, unacknowledged archaisms (many of them--this is the root of the difficulty--firmly engraved, at the very least, in the passive memory of the average literate speaker). The logical avenue of approach to the unadulterated diachronic view is the historical dictionary, provided its materials are so ordered as to bring out plastically the dynamics of lexical development, with heightened attention to the succession and mutual compatibility of meanings. For not a few languages, even those subject to philological inspection, no dictionary aiming uncompromisingly at this particular goal exists or is clearly in sight at present. In the existing historical dictionaries, all too frequently

the meaning listed first is not the oldest on record, nor indeed the oldest by the standards of reconstruction, but the one most familiar to moderns or held "fittest" to have acted as a semantic fountain-head.

Over against the historical ordering of linguistic facts one may place etymological studies, i.e., conjectural research in the prehistory of the lexicon. These inquiries fall into several categories, the simplest criterion of classification being that of scope. Even the full-sized etymological dictionaries are diversified, some being ostensibly confined to a single language (Kluge, Skeat, Wartburg), while others are avowedly comparative (Diez, Meyer-Lübke, Pokorny). In practice one can hardly draw a sharp line of demarcation between isolationists and comparatists; and titles are scarcely reliable guides to treasuries of etymological hypotheses and material. Conceivably the clearest separation lies between those dictionaries which start from the base, real or assumed, and lead the reader to the product(s), and those that invite one to follow the opposite course. At least one recent venture, V. García de Diego's (1954), is bipartite and tries, between two covers, to afford both views in close succession (unfortunately, with less than satisfactory results).

Aside from fully integrated etymological dictionaries one encounters loosely strung miscellanies of etymological notes, of varying length; "gleanings," for the most part provided by gaps and deficiencies in recently published book-length reference works; and all sorts of concomitant references to etymological clues, usually of scant value to the expert--in glossaries attached to anthologies and chrestomathies, in old-fashioned dialect word-lists, and the like. There exist further etymological vocabularies devoted to specific lexical strains--say, Arabisms, Gallicisms, Italianisms, and Lusisms in Spanish; that is to say, closely coherent groups of borrowings.

(b) Three contrasting patterns of arrangement. The basic arrangement of dictionary items may be alphabetic, semantic, or casual. Each of these categories lends itself to further subdivision; a loose combination of the first two patterns is likewise conceivable.

The alphabetic arrangement, though strictly conventional, is so overwhelmingly dominant that the ordinary person associates with this familiar sequence the very genre of the dictionary, on a par with a catalogue, a directory, or a mailing list. Even where an ancient text has been edited with every conceivable paleographic nicety, the accompanying glossary, as a rule, parades the words in normal script, following the alphabetic order. Such occasional deviations and complications as do come to mind represent frills; thus, as regards Old and even Classical Spanish, there exist conflicting traditions for placing words beginning with c, ch, h, i, j, k, y, z, to say nothing of l, n, r vis-à-vis ll, nn (ñ), rr (R). In dialect works the use of phonetic (less so of phonemic) transcription imposes an aberrant sequence of letters on account of special characters and unfamiliar diacritic marks. A few lexicographers assemble in alphabetic order not all words, but only the heads or reputed heads of families. A final formidable obstacle to quick, smooth consultation has been in some countries the inflationary spiral of Supplements and even Supplements to Supplements, a veritable cancer-growth, especially where the accretion amounts to a mere handful of last-minute gleanings.

Remotely akin to the alphabetic dictionary is the rhyming dictionary which ideally absorbs the totality of a lexicon, distributing the items on the basis of form rather than of meaning. What sets it apart is not only the choice of the stressed vowel rather than of the initial letter as the classificatory norm, but also the departure from the straight linear sequence in favor of sundry groups of varying size. Students of suffixes

have derived considerable benefit from rhyming dictionaries and from their close congeners, the "reverse" (rückläufig) dictionaries, of which there exist specimens, e.g., for classical Greek, Latin, and modern Russian.

The semantic (analogical, ideological) dictionary unites words by bonds of meaning. Its roots are old; thus, Isidore of Seville's Etymologiae subdivide the entire material into certain provinces of cosmic structure and human endeavor. Within each group and subgroup so circumscribed, one may, but need not, maintain the alphabetical order. The result of such a compromise is a mixed classification. Consistently semantic ramification is difficult to achieve, because only relatively few and isolated semantic "fields" (e.g., anatomical, color, and kinship terms) fall into a neatly delineated schema. Another prototype of the semantic dictionary is the book-length list of synonyms, actually near-synonyms. The project of such a book, typically—but not exclusively--monolingual, may materialize in different ways (clusters of synonyms supported by well-chosen quotations from authors; "ideological index" to a standard dictionary; array of regional or temporal counterparts of each basic entry--an arrangement sometimes called "onomasiologic" in the Central European tradition of modern-language scholarship.

Concerning recent experiments with the fragmentation of the "world of words" into discrete semantic provinces, one may safely affirm that, where short texts are involved, such an approach has turned out to be unpromising. However, the method, duly refined, may some day yield satisfactory results when extended to distinctly bulkier objects—say, the complete works of a prolific writer noted for his abundant and delicately nuanced lexicon. The worthwhileness of semantic classification, in other words, seems to increase proportionately with the absolute size and the relative density of the network. Meanwhile, Roget's Thesaurus, Boissière's Dictionnaire analogique,

and Casares' Diccionario ideológico remain interesting testing grounds.

If it is true that the time-honored "parts of speech" (as against the modern descriptivist's form classes) were semantically rather than behaviorally defined, then some older dictionaries resorting to them (e.g. P. de Alcalá's Vocabulista arábigo, which, under each letter, subdivides the entries into verbs, nouns, and adverbs, then reserves a concluding cross-alphabetic section for numerals) show close affinity to the semantic ordering.

It is unusual for any book-length dictionary to present its entries in completely haphazard fashion, yet a few books, not only by amateurs, illustrate just this bent for eccentricity, which, the sober-minded may argue, squarely defeats the very purpose of a reference work. A grammar, whether historical or descriptive, even though not primarily planned as a lexicographic guide, may incidentally be so used, in default of a more appropriate source if equipped with a handy index. Without this device the lexically oriented scholar may find its consultation forbiddingly time-consuming--just how prohibitive depends on the neatness and transparency or organization, also on the reader's skill, knack for hunches, and sheer luck. In the case of a loose collection of notes, with etymological, phraseological, or puristic overtones, the alphabetic order is common and desirable, but far from obligatory, especially in periodicals. One also comes across miscellanies displaying a fundamentally free sequence with intermittent stretches of alphabetic ordering.

(c) Three contrasting levels of tone. The tone of a dictionary may be detached, preceptive, or facetious. Only the first possibility insures a rigorously erudite approach; what complicates matters is that some normative and jocose dictionaries, though plainly falling short of the minimum standard for scholarly performance, nevertheless may contain slivers of information which, after proper filtering, are apt to be of real help to the serene researcher.

The category of dictionaries designed to report facts objectively includes the great majority of mono- and bi-lingual reference works and practically all historically slanted glossaries. Among regional word-lists, however, the consensus of taste is less in evidence. Many of these vocabularies are intended to be descriptive, with a light undercurrent of innocuous pride in local traditions. Ordinarily enlightened amateurs are at their best when native dialect speech represents to them neither a matter of painful "guilt or inferiority complex," nor the pretext for some morbidly passionate commitment, still less a source of cheap entertainment, but a quietly treasured possession, a kind of revered family piece surrounded by affectionate curiosity.

Over against this gratifying attitude one finds all too frequently the purist's preceptive (normative, didactic) approach. Regionalisms are collected for the purpose of warning the group of speakers concerned <u>against</u> their use. Outside the domain of dialectology, there also exist all sorts of corrective and remedial dictionaries exposing impropriety of speech and writing--among the semiliterate, certain peculiar population groups (e.g., immigrants), snobs succumbing willingly to the influence of foreign usage, workers endangered by professional hazards (translators, journalists), and the like. Though the analyst may find himself in basic disagreement with the bulk of this output, he can, with a measure of skill, extract useful information from such lists of "dos" and "don'ts," provided they are not entirely arbitrary.

A playful, teasing attitude toward language may be part of a general cultural pattern and heritage. Inevitably, a sophisticated worker's initial reaction to the brochures and pamphlets, mostly by unknowns or incompetents, containing all sorts of lexical scraps and tidbits is violently unfavorable; but, after calming down, he may discover that these humorously tuned collections are not necessarily worthless, if seen and

assessed in the proper perspective, after very careful sifting.

Can any sets of relationship be established between these three major patterns of arrangement? A historical dictionary may correspond to alphabetic, infrequently to semantic, seldom (if ever) to casual ordering, and its tone is likely to be matter-of-fact rather than admonitory or exhilarating; a synchronic dictionary, unless scientifically designed, need not be hemmed in by comparable restrictions. Alphabetic arrangement goes well with a descriptive, preceptive (curative), or entertaining purpose and matches equally the diachronic and the synchronic perspectives, while casual arrangement and amusement are at all times easily reconcilable. These few examples show the tendential rather than obligatory character of all conjunctures.

6. Classification by presentation

The four salient points to which we agreed to reduce the aggregate of narrower preferences were definition, verbal documentation, graphic illustration, and the presence of special features (e.g., localization or phonetic transcription).

The degree of specificness and of fullness in a lexical definition ("gloss") represents some kind of continuum, ranging from extreme parsimony to profusion of technical detail. However desirable the goal of maximum information, one is led to posit, in two directions, a saturation point, beyond which further accumulation of details threatens to detract from the architectonic value of a tightly built dictionary. One such menace is the infiltration of historical or naturalistic excursuses, in most instances, mere bric-à-brac; even those minute descriptions of household utensiles and trade tools that so conspicuously benefit the Wörter-und-Sachen school of etymology belong, strictly speaking, in a portfolio or a manual of material civilization. The other kind of otiose superabundance stems

from the indiscriminate recording of contextual shades of meaning: the harmoniously structured dictionary must remain a compact depository of steady designations.

In appraising documentation, it is advantageous to separate the literary language (past and/or present) from, on the one hand, living dialect speech, unsupported by texts, and, on the other, such samples of regional speech as are reflected in locally colored literature.

Where literary usage alone is at issue, the specificness of documentation may oscillate between (a) a generously carved-out passage, (b) a shorter segment of that passage, (c) a bare reference to line, quatrain, folio, page, or chapter, and (d) a simple computation of frequency—range plus incidence—at best interspersed with occasional quotations of characteristic words in their immediate environments (e.g., qualifiers). A few dictionaries, gambling on the layman's indifference, identify the author cited, but neither the particular work nor the specific locus, thus obviating all possibility of effective control.

In the rare dialect glossaries displaying the chosen words in any kind of meaningful context, the compiler faces the dilemma between coining suitable model sentences of his own (if he happens to be an uninhibited speaker of that dialect) and citing such utterances of his informants as he has unobtrusively overheard or—a quicker but less felicitous alternative technique—deliberately elicited. Some such glossaries have a comparativistic superstructure.

If the tradition of a national literature has favored the production of texts entirely or partially dialectal (an example of the latter would be the dialogues, as against the narrative sections, of a novel or short story), it is possible to manufacture a regional vocabulary based in its entirety on excerpts from written and, as a rule, published sources, or to combine

field-notes and aural impressions with the testimony of belles-lettres. At this juncture, the support of folklore plays a signal part: proverbs, riddles, nursery rhymes, puns, and idiomatic sayings may be either appended as a separate exhibit or absorbed as an essential ingredient, sometimes specifically announced in title or subtitle.

Though stylized adornments embellished a few of the older vocabularies, the inclusion of functional pictorial material became standard parctice only after 1800. For the purpose of a lexicographic, rather than typographic or bibliophilic, survey one may distinguish between drawings, sketches, and photographs, on the one hand, and charts and maps, on the other..

The drawing or sketch in a dictionary adds that dosage of concreteness which quickens the lay reader's grasp of a scientific definition or an abstract description. The photograph contributes the dual touch of authenticity and plastic suggestiveness. The option between drawing and photograph dwindles frequently to a matter of printing costs and available space; where the budget allows for separate plates, these may be inserted between pages, assembled at the end of the volume, or segregated in an accompanying portfolio. An independent volume containing drawings or photographs of objects rich in lexicographic implications or else an imaginative combination of both media is equally conceivable. Pictures pertinent to linguistic inquiries are normally in black and white, the chromatic scale being regarded more as a potential amenity than as an actual asset; nevertheless, multicolored plates, the trademark of leading encyclopedias, might be useful, especially to the sensorially perceptive etymologist, as a frame of reference for the vernacular nomenclature of minerals, flowers, birds, and garments. A different use of photographic plates may be visualized in connection with painstaking editions of medieval glosses and glossaries.

While iconography, as a rule, plays a subordinate role in lexical

undertakings, it may in exceptional cases assume a commanding position. The German "Duden" method embodies a miniature pictorial encyclopedia: in imitation of the "direct" approach, it bluntly juxtaposes label and image, rather than entry and gloss, in a preëminently semantic ordering of the lexicon, thus short-circuiting wearisome definitions. There exist well-established adaptations in numerous other countries. The affinity of the underlying principle and practice to recent trends of language instruction and to current auditive experiments centered around wire and tape recorders is unquestionable. What ties these endeavors together, harmonizing them with broader cultural trends the world over, is their strong sensory, anti-conceptual proclivity.

Regional dictionaries and monographs accompanied by folding maps or smaller charts are no longer uncommon. The inclusion, in studies of this scope, of a small-scale linguistic atlas remains a desideratum, for the most part unfulfilled; and even where one has been provided, as in T. Navarro's monograph on Puerto Rican, complete integration of the lexical and the cartographic sections proves difficult of achievement.

Special features worthy of mention include: the use of abbreviations or peculiar idiogrammatic classifiers to mark for a given word its grammatical or semantic category, social plane or emotional slant, or else territorial limitation; the extra measure of attention accorded to pronunciation, conducive to the parenthetic use of phonetic transcription or, at least, to the consistent marking of stress (a side-line leads from here to the orthological dictionary as an autonomous genre); the narrow specification of locale, unless implied in the announced scope of the project: a message conveyed in a variety of ways, e.g., by substituting arbitrary numerical symbols for recurrent reference to townships and villages and, where generations are contrasted, even to informants.

7. Special dictionaries.

The classificatory schema here advocated is so flexible as to be readily adaptable to all but an inconsequential percentage of lexical compilations. Any vocabulary offering a fair cross-section of a lexical system, however selective or arbitrarily arranged that cross-section may be, should, at least in theory, be analyzable in terms of distinctive features. A few residual groups of vocabularies, however, which cannot be credited with presenting such a cross-section, must be segregated from the common flock.

The words "vocabulary" and, especially, "dictionary" upon occasion apply quite loosely to any reference work arranged by words or names (collection, index, check-list, catalogue, set of paradigmatic tables, gazetteer, etc.); in most instances, the confusion may initially have been caused by the alphabetic order as the common trait. As a result, the visitor to a library finds on dictionary shelves collections of idioms, literary clichés, proverbs, family and given names, toponyms, fictional personages, topics and topoi, and the like.

The second exceptional group comprises highly specialized vocabularies of trades, crafts, arts, and sciences unrepresentative of the core of the common lexicon. Among these, lists of navigational terms, including terms of shipbuilding, have been in special demand since the 16th century. Agriculture has been another long-standing purveyor of specialized dictionaries. A matter of heightened concern to our own guild is the dictionary of linguistic or grammatical terms. The problem lies in drawing a cogent border-line between straight vocabularies of this kind and, among the older vintage of scientific treatises, a few so clearly patterned and, terminologically, so comprehensive as to provide, at a glance, the needed lexicographic information without, in the process, acquiring the status of a formal dictionary, not even one of the "semantic" variety.

8. Interaction of distinctive features.

Systematic inquiry into the interaction of the features here isolated--aside from the three "perspectives" already viewed from this angle--would in itself require a study of monographic proportions, but a few random examples of seemingly admissible and inadmissible (or, at least, common and uncommon) constellations of such traits are easily supplied. Customarily, there prevails a definite relation between size and tone: dictionaries written in an entertaining vein tend to be slender. Again, historical slant and profusion of pictorial illustrations do not go together: maps, sketches, drawings, and photographs befit the synchronic dictionary, geographically oriented. Graphic documentation and a tone either morosely preceptive or irrepressibly jocose are, for all practical purposes, mutually exclusive. Bi- and pluri-lingual dictionaries are very seldom diachronic, overwhelmingly favor alphabetic arrangement, increasingly shy away from pedantry and jokes; also, they display the barest minimum of graphic illustrations, since the cumulative effect of a gloss (i.e., definition) and a drawing would tend to be tautological.

Footnote

1"Distinctive features in lexicography: a typological approach to dictionaries exemplified by Spanish", Romance Philology, XII: 4 (May, 1959), pp. 366-399; XIII: 2 (November, 1959), pp. 111-155.